

EXHIBIT N

12/11/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Sean Bell

Page 1

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

RICHARD KADREY, et al.,)

)

Individual and)

Representative)

Plaintiffs,)

)

v.)

Case No. 3:23-cv-03417-VC

)

META PLATFORMS, INC.,)

)

Defendant.)

)

** HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY **

Videotaped Deposition of SEAN BELL

San Francisco, California

Wednesday, December 11, 2024

Reported Stenographically by

Michael P. Hensley, RDR, CSR No. 14114

DIGITAL EVIDENCE GROUP

1730 M. Street, NW, Suite 812

Washington, D.C. 20036

(202) 232-0646

12/11/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Sean Bell

Page 69

1 the LibGen dataset?

2 A. So before I joined, the -- there was
3 already the decision that LibGen was one of the
4 approved data sources; and so I didn't need to make
5 a new decision to tell an engineer, "Hey, go start
6 collecting LibGen data." It was one of the already
7 existing workstreams.

8 And so, you know, engineers were reporting
9 on the progress of different data they had
10 downloaded or how much a benchmark had gone up or
11 how many tokens had they processed, and so I didn't
12 need to create a new project around it.

13 Q. Okay. And do you know where the engineers
14 who downloaded a copy of LibGen downloaded it from?

15 A. I -- I personally don't know. The -- the
16 sourcing end-to-end of everything that people do is
17 recorded and checked by the auditors, but I don't
18 have firsthand knowledge myself of it.

19 Q. Do you have general knowledge of where the
20 LibGen dataset came from?

21 ATTORNEY HARTNETT: Object to the form.

22 Asked and answered.

12/11/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Sean Bell

Page 70

1 THE WITNESS: I know that's a whole bunch
2 of different sources; so I don't actually know how
3 they got it.

4 BY ATTORNEY STOLER:

5 Q. Have you ever heard of Bibliotik?

6 A. Bibliotik? No.

7 Q. And at the time that you joined the team,
8 was Anna's Archive also an approved dataset for
9 training Llama 4?

10 A. I --

11 ATTORNEY HARTNETT: Object to the form.

12 THE WITNESS: I don't think we had an
13 explicit decision one way or the other.

14 And so Anna's Archive was data -- was a
15 superset of LibGen. And so the decision that I had
16 from legal was that LibGen was an approved data
17 source, and that -- and so it didn't -- at the time
18 I joined, I don't recall it contemplating one way or
19 another about Anna's Archive.

20 And so one of the things that we had to do
21 during this year was figure out, you know, the same
22 reasoning that we had -- you know, the -- that

12/11/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Sean Bell

Page 71

1 applied -- that would've accepted LibGen, you know,
2 would that apply to Anna's Archive?

3 And so -- anyway, so we had to figure that
4 out this year.

5 BY ATTORNEY STOLER:

6 Q. And when you say "that would've accepted
7 LibGen," what do you mean by that?

8 A. Well, I just mean whether or not, like,
9 it's approved for the team to use.

10 Q. You previously mentioned that one of the
11 things that your team had to do during this year
12 was:

13 [As Read] Figure out if the same
14 reasoning that we had applied or that
15 would've accepted LibGen would apply to
16 Anna's Archive?

17 Do you recall that?

18 A. That's what I said, yeah. So to clarify
19 it a little bit, I don't know what the reasoning is
20 behind LibGen in terms of the approval.

21 So, in other words, this was an -- when I
22 started the team, there were -- a number of datasets